09/869198
JAB 1463
JC18 Rec'd PCT/PTO 2 1 JUN 2001

INS
A17

VASCULAR ENDOTHELIAL GROWTH FACTOR-X

FIELD OF THE INVENTION

The present invention is concerned with a novel vascular endothelial growth factor (VEGF) herein designated "VEGF-X", and characterisation of the nucleic acid and amino acid sequences of VEGF-X.

BACKGROUND OF THE INVENTION

Angiogenesis involves formation and proliferation of new blood vessels, and is an essential physiological process for normal growth and development of tissues in, for example, embryonic development, tissue regeneration and organ and tissue repair. Angiogenesis also features in the growth of human cancers which require continuous stimulation of blood vessel growth. Abnormal angiogenesis is associated with other diseases such as rheumatoid arthritis psoriasis and diabetic retinopathy.

Capillary vessels consist of endothelial cells which carry the genetic information necessary to proliferate to form capillary networks. Angiogenic molecules which can initiate this process have previously been characterised. A highly selective mitogen for vascular enothelial cells is vascular endothelial growth factor (VEGF) (Ferrara *et al.*, AVascular Endothelial Growth Factor: Basic Biology and Clinical Implications≅. Regulation of angiogenesis, by I.D. Goldberg and E.M. Rosen 1997 Birkhauser Verlag Basle/Switzerland). VEGF is a potent vasoactive protein which is comprised of a glycosylated cationic 46-49 kd dimer having two 24 kd subunits. It is inactivated by sulfhydryl reducing agents and is resistant to acidic pH and to heating and binds to immobilised heparin.

SUBSTITUTE SHEET (RULE 26)

VEGF-A has four different forms of 121, 165, 189 and 206 amino acids respectively due to alternative splicing. VEGF121 and VEGF165 are soluble and are capable of promoting angiogenesis, whereas VEGF189 and VEGF206 are bound to heparin containing proteoglycans in the cell surface. The temporal and spatial expression of VEGF has been correlated with physiological proliferation of the blood vessels (Gajdusek, C.M., and Carbon, S.J., Cell Physiol., 139:570-579, (1989)); McNeil, P.L., Muthukrishnan, L., Warder, E., D=Amore, P.A., J. Cell. Biol., 109:811-822, (1989)). Its high affinity binding sites are localized only on endothelial cells in tissue sections (Jakeman, L.B., *et al.*, Clin. Invest. 89:244-253 (1989)). The growth factor can be isolated from pituitary cells and several tumor cell lines, and has been implicated in some human gliomas (Plate, K.H. Nature 359:845-848, (1992)). The inhibition of VEGF function by anti-VEGF monoclonal antibodies was shown to inhibit tumor growth in immune-deficient mice (Kim, K.J., Nature 362:841-844, (1993)).

VEGF proteins have been described in the following patents and applications all of which are hereby incorporated by reference EP-0,506,477, WO-95/24473, WO-98/28621, WO-90/13649, EP-0,476,983, EP-0,550,296, WO-90/13649, WO-96/26736, WO-96/27007, WO-98/49300, WO-98/36075, WO-98/840124, WO-90/11084, WO-98/24811, WO-98/10071, WO-98/07832, WO-98/02543, WO-97/05250, WO-91/02058, WO-96/39421, WO-96/39515, WO-98/16551.

The present inventors have now identified a further vascular endothelial growth factor, designated herein as "VEGF-X", and the nucleic acid sequence encoding it, which has potentially significant benefits for the treatment of tumours and other conditions mediated by inappropriate angiogenic activity.

DOCID: <WO 9834950A1>

SUMMARY OF THE INVENTION

In the present application, there is provided a novel vascular endothelial growth factor, herein designated "VEGF-X", nucleic acid molecules encoding said growth factor, an expression vector comprising said nucleic acid molecule, a host cell transformed with said vector and compounds which inhibit or enhance angiogenesis. Also provided is the sequence of a CUB domain present in the sequence of VEGF-X which domain itself prevents angiogenesis and which is used to treat diseases associated with inappropriate vascularisation or angiogenesis.

EP 0 774 513 A2

BRIEF DESCRIPTION OF THE FIGURES

The invention may be more clearly understood with reference to the accompanying example, which is purely exemplary, with reference to the accompanying drawings, wherein:

Figure 1: is a DNA sequence identified in the Incyte LifeSeq™ database coding for a novel VEGF-X protein.

Figure 2: is an illustration of amino acid sequence of the nucleic acid sequence of Figure 1.

Figure 3: is an illustration of PCR primer sequences utilised to identify the VEGF-X protein according to the invention.

Figure 4: is a diagrammatic illustration of the spatial relationships in the VEGF-X sequence of the clones identified using the PCR primer sequences of Figure 3.

Figure 5: is an illustration of the nucleotide sequences of the 5= RACE primers used to identify the 5= end of the VEGF-X open reading frame.

Figure 6: is an illustration of the sequence obtained from the RACE experiment.

Figure 7: is an illustration of the nucleotide sequences obtained from the search of LifeSeq™ database using the sequence in Figure 6.

Figure 8: is an illustration of the primers used to clone the entire coding sequence of VEGF-X.

Figure 9: is an illustration of the entire coding sequence of VEGF-X.

Figure 10: is an illustration of the predicted amino acid sequence of the nucleotide sequence of Figure 9.

Figure 11: is an alignment of the sequence of Figure 10 with the sequences of VEGF-A to D.

Figure 12: is an illustration of variant sequences of the VEGF-X protein according to the invention.

Figure 13: is an illustration of the oligonucleotide primers used for E.coli expression of VEGF-X domains and for expression of the full length sequence

0960-0450

of VEGF-X in a baculovirus/insect cell expression system.

Figure 14: depicts nucleic acid sequences of 18 human EST clones obtained from a BLAST search of the LifeSeq™ database used to identify the full sequence encoding VEGF-X.

Figure 15: depicts the nucleotide sequences of 50 human EST clones obtained from the LifeSeq™ database.

Figure 16: is an illustration of nucleotide sequences utilised as primers to identify the nucleotide sequence encoding VEGF-X.

Figure 17: is a nucleotide sequence coding for a partial VEGF-X protein according to the invention.

Figure 18: is an illustration of a partial nucleotide sequence encoding VEGF-X protein according to the invention.

Figure 19: is an illustration of a DNA and polypeptide sequence used for mammalian cell expression of VEGF-X. The predicted VEGF-X signal sequence is in lower case letters. The C-terminal V5 epitope and His6 sequences are underlined.

Figure 20: is an illustration of a DNA and polypeptide sequence used for baculovirus/insect cell expression of

WO 00/24760 PCT/US99/24760

VEGF-X. In the polypeptide sequence the signal sequence is shown in lower case. The N-terminal peptide tag added to the predicted mature VEGF-X sequence is underlined.

Figure 21: is an illustration of a DNA and polypeptide sequence used for *E. coli* expression of VEGF-X. The polypeptide sequences at the NB and C- termini derived from the MBP fusion and His6 tag respectively are underlined.

Figure 22: illustrates the disulphide-linked dimerisation of VEGF-X. Protein samples were analysed by SDS-PAGE. Prior to loading the gel, samples were heated to 95EC for 5 minutes in sample buffer in the presence (+) or absence (!) of reducing agent. (A) samples from COS cell expression of a C-terminally V5/His6 peptide-tagged construct. The left hand panel is total conditioned medium, the right hand panel is material purified on Nickel agarose resin. Reduced monomer and putative disulphide-linked, non-reduced dimer are indicated by arrows. There appears to be proteolysis of the protein during purification. Gels were blotted onto nylon membranes and protein detected with an anti V5 monoclonal antibody. (B) Samples from *E.coli* expression of a maltose-binding protein/His6 dual fusion construct. M indicates the molecular weight markers (Benchmark, LifeTechnologies). The gel

was stained with Coomassie Blue by standard procedures. The fusion protein has an apparent molecular weight of 80kDa.

Figure 23: illustrates the glycosylation of VEGF-X. VEGF-X was purified from the culture supernatant of COS cells transfected with the pcDNA6/V5-His construct. Supernatants were harvested 72h post-transfection and purified on nickel resin. Samples were then treated with EndoH (+) or untreated (-) before SDS-PAGE and blotting, as described in the legend to Figure 22.

Figure 24: is an illustration of the DNA and polypeptide sequence used for *E. coli* expression of the VEGF-like domain of VEGF-X. Polypeptide sequences at the N-terminus of the protein derived from the vector are underlined.

Figure 25: shows expression of the VEGF-X VEGF domain in *E. coli*. Lane 1-10μl broad range marker (New England Biolabs), lane 2-10μl unreduced sample, lane 3-10μl reduced sample. The reduced PDGF domain protein (lane 3) has an apparent molecular weight of approximately 19kDa on SDS-PAGE.

Figure 26: illustrates a DNA and polypeptide sequence used for *E. coli* expression of the CUB-like domain of VEGF-X. The polypeptide sequence at the N-terminus derived from the vector-encoded signal

and the introduced His6 tag are underlined.

Figure 27: shows expression of the VEGF-X CUB domain in *E. coli*. The CUB domain protein was purified on Nickel chelate resin. The protein migrates at approximately 23kDa on SDS-PAGE.

Figure 28: illustrates the effect of truncated VEGF-X (CUB domain) on HUVEC proliferation. (A) Human Umbilical Vein Endothelial Cells (one-day-treatment). (B) Human Umbilical Vein Endothelial Cells (24-hour starving followed by one-day-treatment). (C) Effect of VEGF-A_{165} and VEGF-X CUB domain on the proliferation of HUVEC (two-day-treatment).

Figure 29: depicts the tissue distribution of VEGF-X mRNA analysed by Northern blotting and RT-PCR in (A) normal tissues and (B) tumour tissue and cell lines.

Figure 30: depicts the partial intron/exon structure of the VEGF-X gene. (A) Genomic DNA sequences of 2 exons determined by sequencing; exon sequence is in upper case, intron sequence is in lower case. (B) Shows the location of splice sites within the VEGF-X cDNA sequence. The location of mRNA splicing events is indicated by vertical lines. The cryptic splice donor/acceptor site at nt. 998/999

(diagonal lines) gives rise to the splice variant forms of VEGF-X. No splice site information is given for the region shown in italics.

Figure 31: is a graphic representation of the effect of FL-VEGF-X on HuVEC proliferation: (24 hour serum starvation followed by one day treatment).

Figure 32: is a graphic representation of the combined effect of truncated VEGF-X (CUB domain) and human recombinant $VEGF_{165}$ on HuVEC proliferation: (24 hour serum starvation followed by two day treatment).

Figure 33: is a graphic representation of the combined effect of the CUB domain and human recombinant bFGF on HuVEC proliferation: (24 hour serum starvation followed by two day treatment).

Figure 34: is a graphic representation of the results of a LDH assay for testing cytotoxicity of the CUB domain or the CUB domain with $rhVEGF_{165}$.

Figure 35: is a graphic representation of the results obtained from a LDH assay for testing cytotoxicity of the CUB domain or CUB domain with rh-bFGF.

WO 00/47750 PCT/US00/03450

DETAILED DESCRIPTION OF THE INVENTION

Therefore, according to a first aspect of the present invention there is provided a nucleic acid molecule encoding a VEGF-X protein or a functional equivalent, fragment, derivative or bioprecursor thereof, said protein comprising the amino acid sequence from position 23 to 345 of the amino acid sequence illustrated in Figure 10. Alternatively, the nucleic acid molecule of the invention encodes the complete sequence identified in Figure 10 and which advantageously includes a signal peptide to express said protein extracellularly. Preferably, the nucleic acid molecule is a DNA and even more preferably a cDNA molecule. Preferably, the nucleic acid molecule comprises the nucleotide sequence from position 257 to 1291 of the nucleotide sequence illustrated in Figure 9. In a preferred embodiment the nucleic acid is of mammalian origin and even more preferably of human origin.

SUBSTITUTE SHEET (RULE 26)

In accordance with the present invention a functional equivalent should be taken to mean a protein, or a sequence of amino acids that have similar function to the VEGF-X protein of the invention.

Also provided by this aspect of the present invention is a nucleic acid molecule such as an antisense molecule capable of hybridising to the nucleic acid molecules according to the invention under high stringency conditions, which conditions would be well known to those skilled in the art.

Stringency of hybridisation as used herein refers to conditions under which polynucleic acids are stable. The stability of hybrids is reflected in the melting temperature (Tm) of the hybrids. Tm can be approximated by the formula:

$$81.5°C+16.6(\log_{10}[Na^{+}]+0.41\ (\%G\&C)-600/l$$

wherein l is the length of the hybrids in nucleotides. Tm decreases approximately by 1-1.5°C with every 1% decrease in sequence homology.

The term Astringency≅ refers to the hybridisation conditions wherein a single-stranded nucleic acid joins with a complementary strand when the purine or pyrimidine bases therein pair with their corresponding base by hydrogen bonding. High stringency conditions favour homologous base pairing whereas low stringency conditions favour non-homologous base pairing.

09891582.012402

"Low stringency" conditions comprise, for example, a temperature of about 37°C or less, a formamide concentration of less than about 50%, and a moderate to low salt (SSC) concentration; or, alternatively, a temperature of about 50EC or less, and a moderate to high salt (SSPE) concentration, for example 1M NaCl.

"High stringency" conditions comprise, for example, a temperature of about 42°C or less, a formamide concentration of less than about 20%, and a low salt (SSC) concentration; or, alternatively, a temperature of about 65°C, or less, and a low salt (SSPE) concentration. For example, high stringency conditions comprise hybridization in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C (Ausubel, F.M. *et al.* Current Protocols in Molecular Biology, Vol. I, 1989; Green Inc. New York, at 2.10.3).

"SSC" comprises a hybridization and wash solution. A stock 20X SSC solution contains 3M sodium chloride, 0.3M sodium citrate, pH 7.0.

"SSPE" comprises a hybridization and wash solution. A 1X SSPE solution contains 180 mM NaCl, 9mM Na_2HPO_4 and 1 mM EDTA, pH 7.4.

The nucleic acid capable of hybridising to nucleic acid molecules according to the invention will generally be at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the nucleotide sequences according to the invention.

The antisense molecule capable of hybridising to the nucleic acid according to the invention may be used as a probe or as a medicament or may be included in a pharmaceutical composition with a pharmaceutically acceptable carrier, diluent or excipient therefor.

The term "homologous" describes the relationship between different nucleic acid molecules or amino acid sequences wherein said sequences or molecules are related by partial identity or similarity at one or more blocks or regions within said molecules or sequences.

The present invention also comprises within its scope proteins or polypeptides encoded by the nucleic acid molecules according to the invention or a functional equivalent, derivative or bioprecursor thereof.

Therefore, according to a further aspect of the present invention, there is provided a VEGF-X protein, or a functional equivalent, derivative or bioprecursor thereof, comprising an amino acid sequence from position 23 to 345 of the sequence as illustrated in Figure 10, or alternatively which amino acid sequence comprises the complete sequence of Figure 10. A further aspect of the invention comprises a VEGF-X

protein, or a functional equivalent, derivative or bioprecusor thereof, encoded by a nucleic acid molecule according to the invention. Preferably, the VEGF-X protein encoded by said nucleic acid molecule comprises the sequence from position 23 to 345 of the amino acid sequence as illustrated in Figure 10, or which sequence alternatively comprises the sequence of amino acids of Figure 10.

The DNA molecules according to the invention may, advantageously, be included in a suitable expression vector to express VEGF-X encoded therefrom in a suitable host. Incorporation of cloned DNA into a suitable expression vector for subsequent transformation of said cell and subsequent selection of the transformed cells is well known to those skilled in the art as provided in Sambrook *et al.* (1989), molecular cloning, a laboratory manual, Cold Spring Harbour Laboratory Press.

An expression vector according to the invention includes a vector having a nucleic acid according to the invention operably linked to regulatory sequences, such as promoter regions, that are capable of effecting expression of said DNA fragments. The term "operably linked" refers to a juxta position wherein the components described are in a relationship permitting them to function in their intended manner. Such vectors may be transformed into a suitable host cell to provide for expression of a polypeptide according to the invention. Thus, in a further aspect, the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell, transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the

polypeptides, and recovering the expressed polypeptides.

The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, and optionally a promoter for the expression of said nucleotide and optionally a regulator of the promoter.

The vectors may contain one or more selectable markers, such as, for example, ampicillin resistance.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for translation initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

Nucleic acid molecules according to the invention may be inserted into the vectors described in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense nucleic acids may be produced by synthetic means.

In accordance with the present invention, a defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including in particular, substitutions in cases which result in a synonymous codon (a different codon specifying the

same amino acid residue) due to the degenerate code in conservative amino acid substitutions. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

The present invention also advantageously provides nucleic acid sequences of at least approximately 10 contiguous nucleotides of a nucleic acid according to the invention and preferably from 10 to 50 nucleotides even more preferably, the nucleic acid sequence comprise the sequences illustrated in Figure 3. These sequences may, advantageously be used as probes or primers to initiate replication, or the like. Such nucleic acid sequences may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be used in diagnostic kits or the like for detecting the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with the sample under hybridising conditions and detecting for the presence of any duplex or triplex formation between the probe and any nucleic acid in the sample.

The nucleic acid sequences according to this aspect of the present invention comprise the sequences of nucleotides illustrated in Figures 3 and 5.

According to the present invention these probes may be anchored to a solid support. Preferably, they are present on an array so that multiple probes can simultaneously hybridize to a single biological sample. The probes can be spotted onto the array or synthesised *in situ* on the array. (See Lockhart *et al.*, Nature Biotechnology, vol. 14, December 1996 "Expression monitoring by hybridisation to high

density oligonucleotide arrays". A single array can contain more than 100, 500 or even 1,000 different probes in discrete locations.

The nucleic acid sequences, according to the invention may be produced using such recombinant or synthetic means, such as for example using PCR cloning mechanisms which generally involve making a pair of primers, which may be from approximately 10 to 50 nucleotides to a region of the gene which is desired to be cloned, bringing the primers into contact with mRNA, cDNA, or genomic DNA from a human cell, performing a polymerase chain reaction under conditions which brings about amplification of the desired region, isolating the amplified region or fragment and recovering the amplified DNA. Generally, such techniques are well known in the art, such as described in Sambrook *et al.* (Molecular Cloning: a Laboratory Manual, 1989).

The nucleic acids or oligonucleotides according to the invention may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S, enzyme labels or other protein labels such as biotin or fluorescent markers. Such labels may be added to the nucleic acids or oligonucleotides of the invention and may be detected using known techniques *per se*.

Advantageously, human allelic variants or polymorphisms of the DNA molecule according to the invention may be identified by, for example, probing cDNA or genomic libraries from a range of individuals, for example, from different populations. Furthermore, nucleic acids and probes according to the invention may be used to sequence genomic DNA from patients using techniques well known in the art, such as the Sanger Dideoxy chain termination method, which may,

advantageously, ascertain any predisposition of a patient to certain disorders associated with a growth factor according to the invention.

The protein according to the invention includes all possible amino acid variants encoded by the nucleic acid molecule according to the invention including a polypeptide encoded by said molecule and having conservative amino acid changes. Conservative amino acid substitution refers to a replacement of one or more amino acids in a protein as identified in Table 1. Proteins or polypeptides according to the invention further include variants of such sequences, including naturally occurring allelic variants which are substantially homologous to said proteins or polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, preferably 80 or 90% and preferably 95% amino acid homology with the proteins or polypeptides encoded by the nucleic acid molecules according to the invention. The protein according to the invention may be recombinant, synthetic or naturally occurring, but is preferably recombinant.

The nucleic acid or protein according to the invention may be used as a medicament or in the preparation of a medicament for treating cancer or other diseases or conditions associated with expression of VEGF-X protein.

Advantageously, the nucleic acid molecule or the protein according to the invention may be provided in a pharmaceutical composition together with a pharmacologically acceptable carrier, diluent or excipient therefor.

The present invention is further directed to

inhibiting VEGF-X *in vivo* by the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation of antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5′ coding portion or the mature DNA sequence, which encodes for the protein of the present invention, is used to design an antisense RNA oligonucleotide of from 10 to 50 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple-helix - see Lee *et al.* Nucl. Acids Res., 6:3073 (1979); Cooney *et al.*, Science, 241:456 (1988); and Dervan *et al.*, Science, 251: 1360 (1991), thereby preventing transcription and the production of VEGF-X. The antisense RNA oligonucleotide hybridises to the mRNA *in vivo* and blocks translation of an mRNA molecule into the VEGF-X protein (antisense - Okano, J. Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)).

Alternatively, the oligonucleotide described above can be delivered to cells by procedures in the art such that the anti-sense RNA and DNA may be expressed *in vivo* to inhibit production of VEGF-X in the manner described above.

Antisense constructs to VEGF-X, therefore, may inhibit the angiogenic activity of VEGF-X and prevent the further growth of or even regress solid tumours, since angiogenesis and neovascularization are essential steps in solid tumour growth. These antisense constructs may also be used to treat rheumatoid arthritis, psoriasis and diabetic retinopathy which are all characterized by abnormal angiogenesis.

A further aspect of the invention provides a host cell or organism, transformed or transfected with an expression vector according to the invention. The host cell or organism may advantageously be used in a method of producing VEGF-X, which comprises recovering any expressed VEGF-X from the host or organism transformed or transfected with the expression vector.

According to a further aspect of the invention there is also provided a transgenic cell, tissue or organism comprising a transgene capable of expressing VEGF-X protein according to the invention. The term "transgene capable of expression" as used herein means a suitable nucleic acid sequence which leads to expression of VEGF-X or proteins having the same function and/or activity. The transgene, may include, for example, genomic nucleic acid isolated from human cells or synthetic nucleic acid, including DNA integrated into the genome or in an extrachromosomal state. Preferably, the transgene comprises the nucleic acid sequence encoding the proteins according to the invention as described herein, or a functional fragment of said nucleic acid. A functional fragment of said nucleic acid should be taken to mean a fragment of the gene comprising said nucleic acid coding for the proteins according to the invention or a functional equivalent, derivative or a non-functional derivative such as a dominant negative mutant, or bioprecursor of said proteins. For example, it would be readily apparent to persons skilled in the art that nucleotide substitutions or deletions may be used using routine techniques, which do not affect the protein sequence encoded by said nucleic acid, or which encode a functional protein according to the invention.

VEGF-X protein expressed by said transgenic cell,

tissue or organism or a functional equivalent or bioprecursor of said protein also forms part of the present invention.

Antibodies to the protein or polypeptide of the present invention may, advantageously, be prepared by techniques which are known in the art. For example, polyclonal antibodies may be prepared by inoculating a host animal, such as a mouse or rabbit, with the polypeptide according to the invention or an epitope thereof and recovering immune serum. Monoclonal antibodies may be prepared according to known techniques such as described by Kohler R. and Milstein C., Nature (1975) 256, 495-497. Advantageously, such antibodies may be included in a kit for identifying VEGF-X in a sample, together with means for contacting the antibody with the sample.

Advantageously, the antibody according to the invention may also be used as a medicament or in the preparation of a medicament for treating tumours or other diseases associated with expression of VEGF-X. The invention also further provides a pharmaceutical composition comprising said antibody together with a pharmaceutically acceptable carrier diluent or excipient therefor.

Proteins which interact with the polypeptide of the invention may be identified by investigating protein-interactions using the two-hybrid vector system first proposed by Chien *et al.*, (1991) Proc. Natl. Acad. Sci. USA 88 : 9578-9582.

This technique is based on functional reconstitution *in vivo* of a transcription factor which activates a reporter gene. More particularly the technique comprises providing an appropriate host cell with a

DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain, expressing in the host cell a first hybrid DNA sequence encoding a first fusion of a fragment or all of a nucleic acid sequence according to the invention and either said DNA binding domain or said activating domain of the transcription factor, expressing in the host at least one second hybrid DNA sequence, such as a library or the like, encoding putative binding proteins to be investigated together with the DNA binding or activating domain of the transcription factor which is not incorporated in the first fusion; detecting any binding of the proteins to be investigated with a protein according to the invention by detecting for the presence of any reporter gene product in the host cell; optionally isolating second hybrid DNA sequences encoding the binding protein.

An example of such a technique utilises the GAL4 protein in yeast. GAL4 is a transcriptional activator of galactose metabolism in yeast and has a separate domain for binding to activators upstream of the galactose metabolising genes as well as a protein binding domain. Nucleotide vectors may be constructed, one of which comprises the nucleotide residues encoding the DNA binding domain of GAL4. These binding domain residues may be fused to a known protein encoding sequence, such as for example, the nucleic acids according to the invention. The other vector comprises the residues encoding the protein binding domain of GAL4. These residues are fused to residues encoding a test protein. Any interaction between polypeptides encoded by the nucleic acid according to the invention and the protein to be tested leads to transcriptional activation of a reporter molecule in a GAL-4 transcription deficient

yeast cell into which the vectors have been transformed. Preferably, a reporter molecule such as β-galactosidase is activated upon restoration of transcription of the yeast galactose metabolism genes.

A further aspect of the present invention also provides a method of identifying VEGF-X in a sample, which method comprises contacting said sample with an antibody according to the invention and monitoring for any binding of any proteins to said antibody. A kit for identifying the presence of VEGF-X in a sample is also provided comprising an antibody according to the invention and means for contacting said antibody with said sample.

SUBSTITUTE SHEET (RULE 26)

VEGF-X may be recovered and purified from recombinant cell cultures by methods known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxyapatite chromatography and lectin chromatography.

The VEGF-X protein of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated with mammalian or other eukaryotic carbohydrates or may be non-glycosylated.

VEGF-X is particularly advantageous as a wound healing agent, where, for example, it is necessary to re-

vascularize damaged tissues, or where new capillary angiogenesis is important. Accordingly, VEGF-X may be used for treatment of various types of wounds such as for example, dermal ulcers, including pressure sores, venous ulcers, and diabetic ulcers. In addition, it can be used in the treatment of full-thickness burns and injuries where angiogenesis is desired to prepare the burn in injured sites for a skin graft and flap. In this case, VEGF-X or the nucleic acid encoding it may be applied directly to the wound. VEGF-X may be used in plastic surgery when reconstruction is required following a burn, other trauma, or even for cosmetic purposes.

09859158.012402

An important application of VEGF-X is to induce the growth of damaged bone, periodontium or ligament tissue. For example, it may be used in periodontal disease where VEGF-X is applied to the roots of the diseased teeth, leading to the formation of new bone and cementum with collagen fibre ingrowths. It can be used for regenerating supporting tissues of teeth, including alveolar bone, cementum and periodontal ligament, that have been damaged by disease and trauma.

Since angiogenesis is important in keeping wounds clean and non-infected, VEGF-X may be used in association with surgery and following the repair of cuts. It should be particularly useful in the treatment of abdominal wounds where there is a high risk of infection.

VEGF-X can also be used for the promotion of endothelialization in vascular graft surgery. In the case of vascular grafts using either transplanted or synthetic material, VEGF-X may be applied to the surface of the graft or at the junction to promote the

growth of the vascular endothelial cells. One derivation of this is that VEGF-X can be used to repair the damage of myocardial and other occasions where coronary bypass surgery is needed by stimulating the growth of the transplanted tissue. Related to this is the use of VEGFX to repair the cardiac vascular system after ischemia.

The protein of the present invention may also be employed in accordance with the present invention by expression of such protein *in vivo*, which is often referred to as "gene therapy".

Thus, for example, cells such as bone marrow cells may be engineered with a polynucleotide (DNA or RNA) encoding for the protein *ex vivo* as defined herein, the engineered cells are then provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, cells may be engineered by procedures known in the art by use of a retroviral particle containing RNA encoding for the protein of the present invention.

Similarly, cells may be engineered *in vivo* for expression of the protein *in vivo*, for example, by procedures known in the art.

A further aspect of the invention comprises a method of treating a disorder mediated by expression of a protein according to the invention, by administering to a patient an amount of an antisense molecule as described herein, in sufficient concentration to alleviate or reduce the symptoms of said disorder.

Compounds which inhibit or enhance angiogenesis may be identified by providing a host cell or organism according to the invention or a transgenic cell,

tissue or organism according to the invention, contacting a test compound with said cell, tissue or organism and monitoring for the effect of said compound compared to a cell tissue or organism which has not been contacted with said compound. These compounds may themselves be used as a medicament or included in a pharmaceutical composition for treatment of disorders mediated by inappropriate vascularisation or angiogenic activity.

The present inventors have also, advantageously, identified in the sequence encoding the VEGF-X protein a CUB domain, which has heretofore not previously been identified in VEGF-type growth factors. The VEGF-X protein may therefore exert dual regulatory effects via interaction with the VEGF tyrosine kinase receptors or with neuropilin receptors mediated by the CUB domain. Thus, the sequence encoding said CUB domain may be included in an expression vector for subsequent transformation of a host cell, tissue or organism.

0969136 .0124462

VEGF-X or fragments thereof may be able to modulate the effects of pro-angiogenic growth factors such as VEGF as indicated in the findings presented in the examples below that the N-terminal part of the VEGF-X protein, a CUB-like domain, is able to inhibit VEGF-stimulated proliferation of HUVECs. VEGF-X or fragments thereof may therefore be useful in therapy of conditions involving inappropriate angiogenesis. Inhibition of the angiogenic activity of VEGF has been linked with inhibition of tumour growth in several models eg Kim K. J. et al, Nature 362:841-844, (1993). Additionally, agents able to inhibit angiogenesis would be expected to be useful in treating other angiogenesis-dependent diseases such a retinopathy, osteoarthritis and psoriasis(Folkman, J., Nature

Medicine 1:27-31, (1995).

As identified in more detail in the Examples described herein the present inventors have surprisingly identified that the CUB domain of VEGF-X is able to inhibit stimulation of proliferation of HUVECs induced by either VEGF or bFGF. The CUB domain may, therefore, be utilised as a therapuetic agent for inhibition of angiogenesis and for treatment of condition associated with inappropriate vascularisation or angiogenesis.

Therefore according to a further aspect of the invention there is provided a method of inhibiting angiogenic activity and inappropriate vascularisation including formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject said method comprising administering to said subject an amount of a polypeptide having an amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 or a nucleic acid molecule encoding the CUB domain according to the invention in sufficient concentration to reduce or prevent said angiogenic activity.

Furthermore there is also provided a method of treating or preventing any of cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy, said method comprising administering to said subject an amount of a polypeptide having an amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 or a nucleic acid molecule encoding the CUB domain according to the invention in sufficient concentration to treat or prevent said disorders.

The CUB domain may also be used to identify compounds that inhibit or enhance angiogenic activity such as

WO 96/39426 PCT/GB96/01260

inappropriate vascularisation, in a method comprising contacting a cell expressing a VEGF receptor and/or a neuropilin 1 or 2 type receptor with said compound in the presence of a VEGF-X protein according to the invention and monitoring for the effect of said compound or said cell when compared to a cell which has not been contacted with said compound. Such compounds may then be used as appropriate to prevent or inhibit angiogenic activity to treat the disorders or conditions described herein, or in a pharmaceutical composition. An antibody to said CUB domain may also be useful in identifying other proteins having said sequences.

The term "therapeutically effective amount" as used herein is an amount that is effective to prevent or alleviate in some way a condition. In general a dosage is employed capable of establishing in the tissue which is to be the subject of the therapy a dose that is beneficial but not toxic.

09869158 .012402

Deposited Plasmids

	Date of Deposit	Accession No.
Plasmid VEGFX/pCR2.1 1TOPO FL	1 March 1999	LMBP 3925
Plasmid VEGFX/pRSETB BD amino acids G230-G345	1 March 1999	LMBP 3926
Plasmid VEGFX/pcR.2.1 FL Clone 9	20 October 1999	LMBP 3977
Plasmid VEGF-X CUB PET22b	20 December 1999	------

The above plasmids were deposited at the Belgian Coordinated Collections of Microorganisms (BCCM) at Laboratorium Voor Moleculaire Biologie-Plasmidencollectie (LMBP) B-9000, Ghent, Belgium, in accordance with the provisions of the Budapest Treaty of 28 April 1977.

A BLAST (Basic Local Alignment Search Tool; Altschul *et al.*, 1990 J. Mol. Biol. 215, 403-410) search was performed in the proprietary LifeSeq™ human EST database (Incyte Pharmaceuticals, Inc., Palo Alto, CA, USA). BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologues. While it is useful for matches which do not contain gaps, it is inappropriate for performing motif-style searching. The fundamental unit of BLAST algorithm output is the

High-scoring Segment Pair (HSP).

Eighteen human EST clones (Figure 14) with high similarity to the previously identified VEGF proteins were identified and a further fifty EST clones (Figure 15) were identified using these sequences as query sequences, allowing us to deduce the putative sequence for the new VEGF-X protein. The sequences obtained were compared to known sequences to determine regions of homology and to identify the sequence as a novel VEGF-type protein. Using the DNA sequence information in the databases we were able to prepare suitable primers having the sequences of VEGF-X 1-10 illustrated in Figure 3 for use in subsequent RACE experiments to obtain the complete DNA sequence for the VEGF-X gene.

CONFIDENTIAL

Cloning

A profile was developed based on the VEGF-like domain in existing VEGF sequences (VEGF-A, B, C and D). This was used to search the public databases and the Incyte LifeSeq™ database. No significant novel matching sequences were found in the public databases. All of the matching sequences found in the LifeSeq™ database (~1000) were assembled to give a smaller number of sequences (~30), which included the known VEGFs and a potential novel VEGF (Figures 1 and 2). This sequence was named VEGF-X.

Oligonucleotides were designed to amplify the VEGF-X sequence from cDNA (Figure 3). The ESTs found in LifeSeq™ were from a range of tissues, with a slight predominance of sequences from ovary, testis, placenta and lung (Figure 14 and 15). Accordingly the oligonucleotides were used to amplify cDNA derived from lung and placenta. First-round PCR products were found at ~200bp larger than the expected sizes, while

3 major species appeared after a second round of PCR amplification, the smallest of which was of the expected size. These fragments were cloned and sequenced. The smallest fragment did indeed have the sequence originally identified from the LifeSeq database, while the others contained insertions (Figure 4).

As the first round of amplification suggested that the major species found in cDNA from ovary and placenta was not that originally identified in the LifeSeq™ database, the focus of effort was switched to the presumed major species (it seemed likely that clones 57, 25-27 and 2.1kb clones 1-3 in Figure 4 represented the major mRNA species). Conceptual translation of the DNA sequences of these cloned PCR fragments indicated that the complete open reading frame was not present in the clones or in the sequence from LifeSeq™. While all clones contained the same sequence in the region of the translation termination codon, indicating that the end of the open reading frame had been identified, the 5′ end of the open reading frame had not been cloned. 5′ RACE experiments were therefore carried out in order to find the start of the reading frame. PCR primers designed for RACE experiments are shown in Figure 5. RACE PCR products were sequenced directly. Sequence could be obtained from the 3′ end of these RACE products but not from the 5′ end; probably because the products were not cloned and were therefore heterogeneous at the 5′ end. This new sequence was assembled with the existing cloned sequence to give the sequence shown in Figure 6. Searching the LifeSeq™ database with this sequence identifies ESTs which extend the sequence a further 140bp in the 5′ direction and a further 160bp in the 3′ direction (Figure 7). This longer contig was used to design

oligonucleotide primers to amplify the entire coding sequence (these primer sequences are shown in Figure 8). PCR was carried out using primers 5'-1 and vegfX10 (in order to clone a "full-length" cDNA), and with primers 5'-1 and vegfX6 (in order to clone the full coding region, see Figure 3 for sequences of vegfX10 and vegfX6). A number of clones were obtained for the shorter fragment, of which clones 4 and 7 contain no PCR errors (sequence of clones 4 & 7 in Figure 9). A single clone was obtained for the longer fragment (clone 9), but this sequence appears to contain 2 PCR errors.

The predicted polypeptide from these longer contigs is shown in Figure 10. Amino acids 1-22 are predicted to encode a signal sequence (von Heijne, 1986, *Nucleic Acids Res.* **14**, 4683-4690). Figure 11 shows an alignment of the protein sequence with VEGFs A-D. The region homologous to the other VEGFs is located towards the C-terminus of the protein. As the VEGF homology domain is expected to belong to the TGF-beta superfamily of growth factors and to consist of a dimer containing both intra- and intermolecular disulphide bonds, initial alignments focussed on the cysteines. However, mapping of the sequence onto the known x-ray structure of the VEGF-A receptor-binding domain (Muller et al (1997) *Proc. Natl. Acad. Sci USA* **94**, 7192-7197) suggests that the alignment in Figure 11 is plausible, as the extra 4 cysteine residues within the VEGF-homology region of VEGF-X (compared to this region of VEGF-A) correspond to residues which are spatially close in VEGF-A, and may therefore be able to form disulphide bonds.

A search of the PFAM database of protein domains with the full-length polypeptide sequence from Figure 10 identifies two domain consensus sequences within the

polypeptide. The more C-terminal domain is a "VEGF" domain: (the known VEGFs all contain this domain and the structure of this region of VEGF-A is similar to that of PDGF). Additionally towards the N-terminus of the polypeptide there is a CUB domain (amino acids ~40-150). The CUB domain is a 100-110 amino acid extracellular domain found in a number of developmentally-regulated proteins. When the full-length protein is used to search the protein databases using the BLAST 2 algorithm, the scores for matches to CUB domain-containing proteins are more significant than those to the other VEGFs. Interestingly, the most significant matches are to the CUB domains of Neuropilins, and Neuropilin-1 was recently identified as a receptor of one of the VEGF-A isoforms VEGF-A_{165} (Soker et al. (1998) *Cell* **92**, 735-745).

Assuming that the variant sequences isolated by PCR (i.e. the smaller PCR fragments) use the same translation initiation site as the full-length sequence, they would result in production of the variant proteins shown in Figure 12. It may be significant that both of these variant proteins retain the CUB domain and delete all or part of the VEGF-like domain. The production of these variant sequences can be explained by the use of a cryptic splice donor/acceptor site within the VEGF-X sequence (figure 30B, between nt. 998/999): one variant arises by splicing out of the region between nt. 729-998, the other by splicing out of the region between nt. 999-1187.

SUBSTITUTE SHEET (RULE 26)

Expression

Full-length expression constructs

Mammalian cells

Clone 4 containing the full CDS of VEGF-X (see Figure 9), was used to generate constructs for expression of full-length protein. The sequence was amplified by PCR and cloned into the vector pCDNA6/V5-His so as to add a C-terminal V5 epitope tag and His_6 tag. The DNA and polypeptide sequence in this vector is shown in figure 19. Transient expression in COS cells followed by western blotting and detection via an anti-V5 mAb demonstrates the secretion of a protein of ~50K into the medium in transfected cells only (Figure 22A). This construct can also be used to generate VEGF-X expressing stable CHO cell lines.

Baculovirus/Insect-cell expression system

For expression in the baculovirus/insect cell system the DNA encoding the predicted mature VEGF-X polypeptide sequence was fused to a sequence encoding a signal derived from melittin, a secreted insect protein. An N-terminal 6His tag was also added to facilitate purification. The insert was then cloned into the baculovirus expression vector pFASTBAC. The DNA and polypeptide sequence of this construct is shown in Figure 20. Infection of *Trichoplusia ni* Hi5 cells with this recombinant baculovirus results in the secretion of a protein of approximately 45K into the medium (data not shown).

E.coli

The coding region of VEGF-X has been cloned in a variety of ways for expression as a secreted protein in *E.coli*. A particularly useful expression clone carries an N-terminal fusion to the *E.coli* maltose-binding protein (MBP- derived from the

expression vector pMAL-p2, New England Biolabs) and a C-terminal fusion to a 6His tag. The DNA and polypeptide sequence of this vector is shown in Figure 21. Sequential purification of cell fractions on Ni-NTA resin and amylose resin allows the isolation of the expressed protein (see Figure 22B).

09869198.012402

Expression of fragments

VEGF

The VEGF domain of VEGF-X has been expressed in *E.coli*. Similar domains from VEGF-A (Christinger et al. (1996) *PROTEINS: Structure, Function and Genetics* **26**, 353-357), and VEGF-D (Achen et al (1998) *Proc. Natl. Acad. Sci USA* **95**, 548-553) have been shown to be capable of binding to the respective receptors. Expression of these domains was carried out using the bacterium *E.coli*. Additionally, the full-length protein was expressed using the baculovirus/insect cell expression system. The oligonucleotide primers which have been obtained for these experiments are shown in Figure 13. The construct directed expression in the bacterial cytoplasm, and as expected the protein was produced in insoluble form in inclusion bodies (the DNA and polypeptide sequence used for PDGF domain expression is shown in figure 24). Inclusion bodies were washed, solubilized with urea and the protein purified under denaturing conditions, before refolding by dialysis to remove the urea. Soluble protein was obtained, but shows little evidence of the disulphide bond linked dimers seen with material derived from animal cells (Figure 25, compare with Figure 22A & B). It is not clear therefore whether this protein is correctly folded.

CUB

The CUB domain has been expressed as a soluble secreted protein in *E.coli* (Figure 26). The protein was purified by binding to Ni-NTA resin (Figure 27) and assayed for activity on HUVECs in an in-vitro proliferation assay.

Properties of the VEGF-X protein

The transient mammalian cell expression system described above has been used to generate full-length

EP 0 984 070 A2

VEGF-X protein, as shown by antibody detection following Western blotting (see Figure 22A).

Disulphide bond linked dimers

The other members of the PDGF family of growth factors, the PDGFs and VEGFs, all exist as dimers in which two monomers constituting the dimer are linked by interchain disulphide bonds. The x-ray structures of PDGF-BB (Oefner et al, 1992), and VEGF-A (Muller et al, 1997) are known and indicate that at least these two members of the family contain two interchain disulphide bonds. Practically this means that in SDS-PAGE analysis of these growth factors the presence of interchain disulphide bonds is shown by a large decrease in mobility in the absence of reducing agent (ie. the nonreduced dimer migrates more slowly through the gel than the reduced monomer). This effect was also expected for VEGF-X, and has been demonstrated for the material obtained from transient mammalian cell expression (Figure 22A). In the case of the full length material produced in *E.coli* only some 10% of the total VEGF-X protein appears to be present as disulphide bond-linked dimers (figure 22B). However, these results provide evidence that the mammalian cell-derived protein is correctly folded, and that a portion of the *E.coli*-derived protein is too.

09869158.012402

Glycosylation

There are 3 predicted potential N-linked glycosylation sites within the VEGF-X protein: at residues 25, 55 and 254 of the polypeptide sequence. The predicted molecular mass of the mature VEGF-X protein is 40kDa, but SDS-PAGE and western blotting (detection via an introduced C-terminal epitope tag- see Figure 19) of the full-length protein expressed in COS cells gives a band slightly larger than the expected size (45-50kDa) as well as one at 25kDa (Figure 22A). This smaller band is presumed to be a C-terminal proteolysis

fragment derived from the full-length molecule (controls from uninfected cells do not show this band), probably corresponding to a cleavage between the CUB and VEGF domains. EndoH treatment of the preparation gives a slight mobility change for the full-length protein (Figure 23), but for the smaller VEGF domain fragment there is a clear change, indicating that the predicted glycosylation site within the VEGF domain at residue 254 is indeed glycosylated.

Activity of proteins in cell-based assays

Protein samples were tested for activity in cell proliferation, cell migration and *in-vitro* angiogenesis assays. Active samples can also be tested in the *in vivo* matrigel mouse model of angiogenesis.

Full-length VEGF-X protein

Conditioned medium derived from COS cells transiently expressing VEGF-X (see Figure 22A) displayed no detectable activity in any of the assays. However, as VEGF-X protein could only be detected in this preparation by Western blotting, and not by Coomassie-staining of gels, it is clearly present at very low levels and this may be the reason for the observed lack of activity in the cell proliferation, migration or *in vitro* angiogenesis tests.

VEGF domain

The VEGF domain protein described above has been tested in cell proliferation (on a range of cell types), cell migration and in vitro angiogenesis assays and has failed to show activity in any of these tests. As suggested above, this may be due to incorrect folding of this protein.

WO 2004/080480 PCT/GB2004/000950

CUB domain

The CUB domain protein at the highest dose tested (1µg/ml) appears to inhibit proliferation of HUVECs in the absence of other stimulation (figure 28A & B). This effect is also seen following stimulation with the lowest VEGF-A_{165} dose tested (1ng/ml- figure 28C). The CUB domain of VEGF-X therefore appears to show antiproliferative activity on HUVECs, even in the presence of low VEGF-A_{165} doses.

Tissue distribution of mRNA

VEGF-A mRNA expression has been shown to be upregulated in a wide variety of human tumors (lung, breast, ovarian, colon, stomach, liver, pancreas, kidney, bladder and prostate- Takahashi et al, 1995). Tumor VEGF-A expression has been shown to correlate with tumor growth rate, microvascular density and tumor metastasis (Takahashi et al, 1995). It was thus of interest to examine the mRNA expression patterns of VEGF-X. Accordingly, Northern blot analysis of mRNA derived from different tissues has been carried out. The results indicate that although the VEGF-X mRNA is expressed at low levels, it is present in a wide range of tissues. PCR amplification of cDNA from a range of tissue sources supports this idea (Figure 29A). The major mRNA species is approximately 3.1kb in size. There is no significant upregulation seen in tumour cell lines or in tumour tissues tested (Figure 29B), with the possible exception of the cell lines GI-117 (lung carcinoma) and SaOS-2 (osteosarcoma). The results of these initial tissue distribution studies do not, therefore, provide evidence for upregulation of VEGF-X in tumour growth, as is seen with VEGF-A.

WO 98/07880 PCT/US97/15150

Genomic structure of the VEGF-X gene

A genomic BAC clone covering the 3' part of the VEGF-X locus was isolated by hybridisation screening of nylon filters containing a human BAC library. Direct sequencing of this clone using oligonucleotide primers based on the VEGF-X cDNA sequence allowed the determination of several intron/exon boundaries (Figure 30). Interestingly, the position of the mRNA splice site within the PDGF domain (nt 1187/1188 in figure 30B) is conserved with respect to those in the VEGF-A and VEGF-D genes (Tischer et al, 1991; Rocchigiani et al, 1998).

EP 1 038 963 A2

Materials & Methods

PCR, Cloning, DNA sequence determination and BAC screening.

All primers were purchased from Eurogentec, Seraing, Belgium. Insert-specific sequencing primers (15- and 16-mers) were designed by visual inspection of the DNA sequences. DNA was prepared on Qiagen-tip-20 columns or on Qiaquick spin columns (Qiagen GmbH, Düsseldorf, Germany) and recovered from the spin columns in 30µl Tris/EDTA-buffer (10mM TrisHCl pH 7.5, 1 mM EDTA (sodium salt)). Sequencing reactions were performed using BigDye™ Terminator Cycle Sequencing Ready Reaction kits (Perkin Elmer, ABI Division, Foster City, CA, USA) and were run on an Applied Biosystems 377 DNA sequencer (Perkin Elmer, ABI Division, Foster City, CA, USA). Polymerase chain reactions were carried out according to standard procedures (Ausubel et al, 1997). The PCR fragments were cloned into vectors pCR2.1 (Invitrogen, Carlsbad, CA. USA) or pCR-TOPO (Invitrogen,NL) according to the manufacturer's instructions. One of those vectors, plasmid VEGFX/pCR2.1 1TOPO FL was deposited on 1 March 1999 under Accession No. LMBP 3925. After sequence determination, the inserts were

cloned into the desired expression vectors (see Figures 19, 20, 21, 24 & 26).

A human genomic BAC library (Genome Systems, Inc., St Louis, MI, USA) was screened by hybridisation to oligonucleotides derived from the VEGF-X cDNA sequence, according to the manufacturer's instructions. BAC DNA was prepared using a Qiagen plasmid midi kit (Qiagen GmbH, Düsseldorf, Germany) according to the manufacturer's instructions with some modifications (after clearing of the lysate from chromosomal DNA, supernatants from individual preparations were pooled on a single column (tip 100), and after the 70 % EtOH wash, the pellet was resuspended overnight at 41C in 100 µl TE). 20-mer sequencing primers were designed based on the known cDNA sequence, and sequencing carried out as above.

5' RACE

In order to extend the cDNA clone in a 5' direction RACE reactions were carried out. Since it was known that the mRNA is present in placenta and skeletal muscle, Marathon-Ready™ placenta and skeletal muscle cDNAs were purchased from Clontech (Palo Alto CA. USA) and used according to the manufacturer=s instructions. DNA fragments were excised from agarose gels, purified using QiaQuick PCR purification columns (Qiagen GmbH, Düsseldorf, Germany) and sequenced directly.

VEGF-X protein expression and purification

DNA fragments encoding the desired protein sequences were amplified by PCR and cloned into appropriate expression vector systems.

For mammalian cell expression, the full coding sequence was cloned into the vector pcDNA6/V5-his

(Invitrogen Leek, NL, see Figure 19 for construct sequence), so as to add a C-terminal peptide tag to assist in detection and purification.

For insect cell expression the sequence of the predicted mature polypeptide was initially amplified to add an N-terminal 6His peptide and then cloned into the pMelBacB vector (Invitrogen, Leek, NL) to add an insect cell signal sequence. The entire insert was then PCR-cloned into the vector pFASTBAC-1 (LifeTechnologies, Gaithersburg, MA, USA) for construction of a baculovirus according to the manufacturer's instructions.

09886159.012402

For *E.coli* expression, the coding region was PCR amplified to add a C-terminal 6His tag and then cloned into the vector pMAL-p2 (New England Biolabs, Beverly, MA, USA). The coding sequence of this construct is shown in Figure 21). The protein was purified first on Ni-NTA resin (Qiagen GmbH, Düsseldorf, Germany) and then on amylose resin (New England Biolabs, Beverly, MA, USA), according to the manufacturers' instructions.

DNA sequences encoding the CUB and VEGF domain fragments of VEGF-X were PCR amplified and cloned into pET22b and pET21a (Novagen, Madison, WI, USA) respectively. The CUB domain protein was prepared either from the periplasm or medium of induced cultures by standard methods (Ausubel et al, 1997). The protein was initially purified by precipitation with 20% ammonium sulphate. After overnight dialysis vs 20mM Tris Hcl pH7.5, 100mM NaCl to remove ammonium sulphate, the protein was further purified on Ni-NTA resin as described above. The VEGF domain protein was expressed in insoluble form, and preparation of inclusion bodies was carried out using standard

procedures (Ausubel et al 1997). Inclusion bodies were dissolved in 6M guanidine hydrochloride, 20mM Tris Hcl pH8.0, 200mM NaCl, 1mM 2-mercaptoethanol, and purified on Ni-NTA resin (Qiagen GmbH, Düsseldorf, Germany) according to the manufacturer's instructions. The protein was refolded by dialysis against several changes of buffer containing decreasing concentrations of denaturant.

Analysis of protein glycosylation was carried out using EndoH (Roche Molecular Biochemicals, Brussels, BE) according to the manufacturer=s instructions.

Cell Proliferation Assay

Human umbilical vein endothelial cells (HUVECs) (Clonetics, San Diego, CA.) were trypsinized with 0.05% trypsin/0.53mM EDTA (Gibco, Gaithersburg, MD.), resuspended in the EGM-2(Clonetics, San Diego, CA.), counted, and distributed in a 96-well tissue culture plate at 5,000 cells/well. Following cell attachment and monolayer formation (16 hours), cells were stimulated with various concentrations of truncated VEGF-X (CUB domain or VEGF domain) or dilutions of culture supernatants of the full-length VEGF-X (COS 7 or HEK293) in DMEM (Gibco, Gaithersburg, MD.) containing 0.5% to 2% FBS (HyClone, Logan, UT) as indicated. For human fetal dermal fibroblasts (American Type Culture Collection, Rockville, MD.), the growth medium was replaced by DMEM containing 0.1% BSA (Sigma, St. Louise, MO.) with or without various concentrations of truncated VEGF-X proteins. For HCASMC (Clonetics, San Diego, CA.), the medium was replaced by DMEM containing 0.5% FBS. The cells were treated for a further 24 hr-72 hr. For the measurement of proliferation, the culture media were replaced with 100 µl of DMEM containing 5% FBS and 3 µCi/ml of [3H]-thymidine (Amersham, Arlington Heights,

IL.). Following pulse labeling, cells were fixed with methanol/acetic acid (3:1, vol/vol) for 1 hour at room temperature. The cells were washed twice with 250 µl/well of 80% methanol. The cells were solubilized in 0.05% trypsin (100µl/well) for 30 minutes then in 0.5% SDS (100 µl/well) for another 30 minutes. Aliquots of cell lysates (180 µl) were combined with 2 ml of scintillation cocktail (Fisher, Springfiled, NJ) and the radioactivity of cell lysates was measured using a liquid scintillation counter (Wallac 1409). In each case, samples were performed in quadruplicate.

Chemotaxis Assay

The chemotactic response of HUVECs was assayed using a 48-well modified Boyden chamber (NeuroProbe, Cabin John, MD.) and collagen-coated (0.1mg/ml type I collagen, Collaboratic Biomedical, Bedford, MA.) polycarbonate membrane filters with a pore diameter of 8 µm (NeuroProbe, Cabin John, MD.). Cell suspensions (15,000/well) were loaded to the upper part of the chemotaxis chamber and stimulated for 4 hours with $rhVEGF_{165}$ (0.1-10 ng/ml) (Calbiochem, San Diego, CA.) or various concentrations of truncated VEGF-X (PDGF domain). Cells remaining on the top of the membrane were removed. Migration was assessed by counting the number of cells that migrated to the lower side of the filter membrane. The membrane was fixed with 10% formaldehyde for 15 min, followed by staining with Gill's hemotoxylin III (Poly Scientific, Bay Shore, NY.). The assay was performed in triplicates and six independent high power fields per well were counted using a light microscope at 250 magnification. The results were expressed as the fold of unstimulated cells (EGM containing 0.1% BSA).

***In Vitro* Angiogenesis Assay**

In vitro angiogenesis in fibrin gels was quantitated using spheroids of human umbilical vein endothelial cells (Korff et al., 1998). To generate endothelial cell spheroids of defined size and cell number, a specific number of cells (~ 800 cells per spheroid) was suspended in EGM-2 culture medium containing 20% methylcellulose (Sigma, St. Louis, MO.), seeded into nonadherent round-bottom 96-well plates. All suspended cells in one well contributed to the formation of a single endothelial cell spheroid within 24 hours. A fibrin gel stock solution was prepared freshly prior to use by mixing 3mg/ml fibrinogen (Calbiochem, San Diego, CA.) in Medium 199(Gibco, Gaithersburg, MD.). Assays were performed in 24-well culture plates. The 1ml fibrinogen stock was mixed with 50 HUVEC spheroids and the corresponding test substance including rh-$VEGF_{165}$ or various concentration of VEGF-X. The spheriod-containing fibrinogen was rapidly transferred into 24-well plates. Fifteen microliters of thrombin (100 NIH U/ml stock, Sigma, St. Louis, MO.) was added to the gel for the fibrin gel formation. The gel formation usually occurred within 30 seconds. After gel formation, 1ml/well of Medium 199 supplemented with 20% FBS, 1mg/ml _-aminocaproic acid (Calbiochem, San Diego, CA.) and antibiotics were added. The gel was incubated at 37°C (5%CO_2, 95% air, 100% humidity). After 3 days, *in vitro* angiogenesis was quantitated by measuring the length of the three longest capillary sprouts that had grown out of each spheroid (100X magnification), analyzing at least 10 spheroids per experimental group and experiment.

Matrigel Mouse Assay

The matrigel mouse assay is carried out as described by Passanti et al (1992).

Analysis of VEGF-X gene expression by RT-PCR analysis.

Oligonucleotide primers VEGF-E2 and VEGF-X14 (Figure 16; Figure 5) were used for the specific PCR amplification of a 350 bp fragment from VEGF-X. PCR amplifications were performed on human multiple tissue cDNA (MTC™) panels (Clontech human MTC panels I and II and human Tumor MTC panel) normalised to the mRNA expression levels of six different housekeeping genes. In addition, cDNA was made from different tumor cell cultures (Caco-2 colorectal adenocarcinoma; T-84 colorectal carcinoma; MCF-7 breast adenocarcinoma; T-47D breast ductal gland carcinoma; HT1080 bone fibrosarcoma; SaOS-2 osteosarcoma; SK-N-MC neuroblastoma; HepG2 hepatoblastoma; JURKAT T-cell leukemia and THP-1 myelomonocytic leukemia). For the preparation of tumor cell cDNA, cells were homogenised and total RNA prepared using the RNeasy Mini kit (Qiagen GmbH, Hilden, Germany) according to manufacturer's instructions. 1 µg of total RNA was reverse transcribed using oligo(dT)15 as a primer and 50 U of Expand™ Reverse Transcriptase (Boehringer Mannheim, Mannheim, Germany) according to the manufacturer's instructions. PCR reactions with VEGF-X-specific or glyceraldehyde-3-phosphate dehydrogenase (G3PDH)-specific primers were then performed on 1 µl of this cDNA. For all cDNAs, PCR reactions with VEGF-X specific primers were performed in a total volume of 50 µl, containing 5 µl (" 1 ng) of cDNA, 1x Advantage KlenTaq PCR reaction buffer, 0.2 mM dNTP, 250 nM of primers VEGF-E2 and VEGF-X14 and 1 µl of Advantage KlenTaq polymerase mix. Samples were heated to 95°C for 30 s and cycling was done for 30 s at 95°C and 30 s at 68°C for 25, 30 or 35 cycles. Control reactions using specific primers that amplify a 1 kb fragment of the housekeeping gene G3PDH were also performed according to the manufacturer's instructions.

Northern blot analysis of VEGF-X.

Northern blots containing 2 µg of poly(A)-rich RNA derived from different human tissues (Clontech Laboratories; MTN™ blot, MTN™ blot II and Cancer Cell Line MTN™ blot) were hybridized according to the manufacturers instructions with a α-[^{32}P]-dCTP random-priming labelled (Multiprime labelling kit, Roche Diagnostics) 293 bp specific VEGF-X fragment (*Pin*AI-*Stu*I fragment including 92 bp of the 3= end coding region and 201 bp of the 3' untranslated region of VEGF-X). The blots were hybridized overnight at 68°C and final washes at high stringency were at 68°C in 0.1x SSC/0.1 % SDS. The membranes were autoradiographed for 1 to 3 days with intensifying screens.

Full length VEGF-X

The effect of full length VEGF-X on proliferation of HuVEC cells was determined by the ^{3}H-Thymidine incorporation assay. HuVEC cells were serum starved for 24 hours prior to treatment with the full length VEGF-X at the concentration range from 100 pg/ml-10 µg/ml. There was no effect of VEGF-X at 100 pg/ml-10 ng/ml on endothelial cell proliferation. At the higher concentrations of FL-VEGF-X (100 ng/ml and 1 µg/ml) there was a marked inhibition of endothelial cell proliferation. This is probably due to the very high endotoxin level in the samples. The VEGF-X sample was purified in order to decrease the endotoxin level and is currently tested in the cell proliferation assay.

The Summary from Testing the CUB Domain

The effect of CUB domain on inhibition of HuVEC prolieration either serum- (2%), rh-VEGF or bFGF-

stimulated, was assessed by the ^{3}H-Thymidine incorporation assay. Cells were serum starved followed by the treatment with the CUB domain and various growth factors. Results showed that the CUB domain inhibited endothelial cell proliferation, either serum- (2%), rh-VEGF or bFGF-stimulated in a dose dependent manner with maximal inhibition at 10 μg/ml. There was approximately a 2-fold inhibition of proliferation (at 10 μg/ml) of cells stimulated with VEGF and bFGF and nearly a 5-fold inhibition of cells stimulated with serum (2%). Results with the LDH assay showed that there was no cytotoxicity associated with the inhibition of cell proliferation by the CUB domain.

09869158.041202

Therefore, the N-terminus of the polypeptide from Figure 10 has been shown to possess a CUB domain. When database searches are carried out using the full-length coding sequence the best matches (i.e. for a BLAST search, those with the lowest probability score) are found with the CUB domain rather than with the VEGF-like domain. The best match from searching release 37 of the SWISSPROT database (Feb 1999) is to the CUB domain of a neuropilin from Xenopus laevis, and the matches to the CUB domains of human neuropilins 1 and 2 are also more significant than matches to the VEGFs.

This similarity is provocative, given the identification of neuropilin-1 and -2 as cellular receptors for the VEGF-A 165 (Stoker *et al.* 1998, reviewed in Neufeld *et al.* 1999). It is plausible therefore that VEGF-X could exert dual regulatory effects: via interaction with the tyrosine kinase VEGF-receptors mediated by the VEGF-like domain, as well as via interaction with VEGF isoforms or with the neurophilin receptors, mediated by the CUB domain.

To the best of our understanding the latter would be entirely novel, and searches on the most recent release of the Incyte database do not reveal any other proteins containing both CUB and VEGF-like domains. This arrangement of domains suggests possible positive or negative models of regulation:

Positive- the VEGF-like domain is able to interact productively with the tyrosine kinase VEGF receptors giving activation, and the CUB domain is able to interact productively with the neuropilin receptor giving activation.

Negative- the VEGF-like domain does not interact productively with the tyrosine kinase VEGF receptors, either preventing receptor dimerisation or blocking the VEGF binding sites. Further, the CUB domain does not interact productively with the neuropilin receptors, either preventing receptor activation or blocking the VEGF binding sites, or indeed by binding to VEGF isoforms and preventing their interaction with receptors.

WO 00/44782 PCT/GB00/00252

TABLE 1

ORIGINAL RESIDUE	EXEMPLARY SUBSTITUTIONS
ALA	SER, THR
ARG	LYS
ASN	HIS, SER
ASP	GLU, ASN
CYS	SER
GLN	ASN, HIS
GLU	ASP, GLU
GLY	ALA, SER
HIS	ASN, GLN
ILE	LEU, VAL, THR
LEU	ILE, VAL
LYS	ARG, GLN, GLU, THR
MET	LEU, ILE, VAL
PHE	LEU, TYR
SER	THR, ALA, ASN
THR	SER, ALA
TRP	ARG, SER
TYR	PHE
VAL	ILE, LEU ALA
PRO	ALA

Application Page

References

1. Ausubel, FM, R Brent, RE Kingston, DD Moore, JG Seidman, JA Smith, K Struhl (Eds). (1997) *Current Protocols in Molecular Biology*, John Wiley and Sons.
2. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683-4690.
3. Muller, YA, B Li, HW Christinger, JA Wells, BC Cunningham and AM de Vos. (1997) Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site. *Proc. Natl. Acad. Sci USA* **94**, 7192-7197.
4. Korff, T and Augustic, H.G. (1998) Integration of endothelial cells in multicellular spheroids prevents apoptosis and induced differentiation. *The Journal of Cell Biology*. 143, 1341-1352
5. Christinger, HW, YA Muller, LT Berleau, BA Keyt, BC Cunningham, N Ferrara and AM de Vos. (1996) *PROTEINS: Structure, Function and Genetics* **26**, 353-357.
6. Achen, MG, M Jeltsch, E Kukk, T Makinen, A Vitali, AF Wilks, K Alitalo and SA Stacker. (1998) *Proc. Natl. Acad. Sci USA* **95**, 548-553.
7. Siemeister, G, B Schnurr, K Mohrs, C Schachtele, C Marme and G Martiny-Baron. (1996) *Biochem. Biophys. Res. Commun.* **222**, 249-255.
8. Soker, S, S Takashima, HQ Miao, G Neufeld and M Klagsbrun (1998). Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor, *Cell* ***92*****:** 735-745.
9. Neufeld, G, T Cohen, S Gengrinovitch and Z Poltorak (1999). Vascular endothelial growth factor and its receptors, *FASEB J.* **13:**9-22.

10. Oefner, C., D'Arcy, A., Winkler, F.K., Eggimann, B. and Hosang, M. (1992). Crystal structure of human platelet-derived growth factor BB. *EMBO J.* **11**, 3921-3926.

11. Passanti, A.., Taylor, R.M., Pili, R., Guo, Y., Long, P.V., Haney, J.A., Pauly, R., Grant, D.S. and Martin, G.R. (1992) A simple, quantitative method for assessing angiogenesis and antiangiogenic agents using reconstituted basement membrane, heparin and fibroblast growth factor. *Laboratory Investigation*, **67**, 519-528.

12. Rocchigiani, M., Lestingi, M., Luddi, A., Orlandini, M., Franco, B., Rossi, E., Ballabio, A., Zuffardi, O. and Oliviero, S. (1990). Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1 between the PIGA and the GRPR genes. *Genomics*, **47**, 207-216.

13. Takahashi, Y., Kitadai, Y., Bucana, C.D., Cleary, K.R. and Ellis, L.M. (1995). Expression of vascular endothelial growth factor and its receptor, KDR, correlates with vascularity, metastasis and proliferation of human colon cancer. *Cancer Research*, **55:** 3964-3968.

14. Tischer, E., Mitchell, R., Hartman, T., Silva, M., Gospodarowicz, D., Fiddes, J.C. and Abraham, J.A. (1991). The human gene for vascular endothelial growth factor: Multiple protein forms are encoded through alternative exon splicing. *J. Biol. Chem.* **266**, 11947-11954.

SEQUENCE LISTING

Sequence ID No 1 corresponds to the amino acid sequence from position 23 to 345 of the amino acid sequence illustrated in Figure 10.

Sequence ID No 2 is the amino acid sequence illustrated in Figure 10.

Sequence ID No 3 corresponds to the sequence from position 257 to 1291 of the nucleotide sequence illustrated in Figure 9.

Sequence ID No 4 corresponds to the polynucleotide sequence of VEGFX1 illustrated in Figure 3.

Sequence ID No 5 corresponds to the polynucleotide sequence of VEGFX2 illustrated in Figure 3.

Sequence ID No 6 corresponds to the polynucleotide sequence of VEGFX3 illustrated in Figure 3.

Sequence ID No 7 corresponds to the polynucleotide sequence of VEGFX4 illustrated in Figure 3.

Sequence ID No 8 corresponds to the polynucleotide sequence of VEGFX5 illustrated in Figure 3.

Sequence ID No 9 corresponds to the polynucleotide sequence of VEGFX6 illustrated in

Figure 3.

Sequence ID No 10 corresponds to the polynucleotide sequence of VEGFX7 illustrated in Figure 3.

Sequence ID No 11 corresponds to the polynucleotide sequence of VEGFX8 illustrated in Figure 3.

Sequence ID No 12 corresponds to the polynucleotide sequence of VEGFX9 illustrated in Figure 3.

Sequence ID No 13 corresponds to the polynucleotide sequence of VEGFX10 illustrated in Figure 3.

Sequence ID No 14 corresponds to the polynucleotide sequence of VEGFX11 illustrated in Figure 4.

Sequence ID No 15 corresponds to the polynucleotide sequence of VEGFX12 illustrated in Figure 4.

Sequence ID No 16 corresponds to the polynucleotide sequence of VEGFX13 illustrated in Figure 4.

Sequence ID No 17 corresponds to the polynucleotide sequence of VEGFX14 illustrated in Figure 4.

Sequence ID No 18 corresponds to the polynucleotide sequence 5'-1 in Figure 8.

Sequence ID No 19 corresponds to the polynucleotide sequence 5'-2 in Figure 8.

Sequence ID No 20 corresponds to the polynucleotide sequence of VEGFX6 illustrated in Figure 13.

Sequence ID No 21 corresponds to the polynucleotide sequence of VEGFX7 illustrated in Figure 13.

Sequence ID No 22 corresponds to the polynucleotide sequence of VEGFX8 illustrated in Figure 13.

Sequence ID No 23 corresponds to the polynucleotide sequence of VEGFX9 illustrated in Figure 13.

Sequence ID No 24 corresponds to the polynucleotide sequence of VEGBAC1 illustrated in Figure 13.

Sequence ID No 25 corresponds to the polynucleotide sequence of VEGBAC2 illustrated in Figure 13.

Sequence ID No 26 corresponds to a polypeptide having the amino acid sequence from amino acid position 40 to 150 of the sequence of Figure 10.

Sequence ID No 27 corresponds to a polypeptide having the amino acid sequence illustrated in Figure 26.

Sequence ID No 28 corresponds to the sequence from

SUBSTITUTE SHEET (RULE 26)

position 5 to 508 of the nucleotide sequence illustrated in Figure 26.

Sequence ID No 29 corresponds to the nucleotide sequence from position 5 to 508 of the nucleotide sequence illustrated in Figure 26.

Sequence ID No 30 corresponds to the sequence from position 214 to 345 of the nucleotide sequence illustrated in Figure 10.